Supplementary Material

Supplementary Methods

Setting, Subjects, and Specimens

This prospective cohort study was conducted through the general academic pediatrics outpatient clinic at the Ann & Robert H. Lurie Children's Hospital of Chicago and Northwestern University Feinberg School of Medicine (NUFSM). Institutional Review Boards at Lurie Children's and NUFSM approved this study, and mothers of infant subjects provided informed consent. Healthy full-term infants ≤ 2m old without previous admission to a neonatal intensive care unit were enrolled into this longitudinal observational study between September 2014 and January 2017. Breastfeeding and household exposure data were collected by parent interview at study onset and study completion. Parents were requested to provide infant stool samples at each well-child visit between enrollment and their first routine blood draw (approximately age 2m, 4m, 6m, and 9-12m). Additional stool samples were collected at intercurrent sick visits when able. To minimize subject discomfort, serum was collected from infants at age 9-12m concurrent with venipuncture performed for anemia and lead exposure screening as part of routine pediatric preventive care. To be evaluable, subjects were required to provide serum at 9-12m. Umbilical cord blood was collected from 50 consecutive full-term deliveries (these mothers were unrelated to study infants) at Prentice Women's Hospital and serum was extracted for serological studies as described below.

Microbiology

Stools from infants were derived from soiled diapers provided by the parents. Stool was extracted from diapers, aliquoted, and stored at -80^{o}C until ready for batch processing of study-related assays. Thawed infant stool samples underwent several assays to assess for toxigenic (TCD) and non-toxigenic *C. difficile* (NTCD) colonization. Stools underwent both glutamate dehydrogenase (GDH; *C. difficile* common antigen) and toxin A/B enzyme immunoassay (QUIK CHEK Complete; Techlab, Blacksburg, VA), as

well as *tcdB* (toxin B gene) PCR (GeneXpert *C. difficile*, Cepheid, Sunnyvale, CA) testing per manufacturer's instructions. Anaerobic stool culture (including alcohol shock for culture enrichment) was performed, as previously described.[1] A single *C. difficile* colony from culture was randomly selected for whole genome sequencing (WGS; methods detailed below). Genomes were queried for the presence (at least 85% sequence identity) of *tcdA* (GenBank Accession: CAJ67494.1), *tcdB (*GenBank Accession: CAJ67492.1), and binary toxin (*cdtA*/*cdtB; (*GenBank Accession: CBE05856.1/ CBE05858.1) using BLAST.[2] TCD colonization was identified if stools were positive for the following: toxin EIA, *tcdB* PCR, and/or a cultured *C. difficile* isolate whose genome contains *tcdA* or *tcdB*. NTCD colonization was identified if stools were GDH-positive but toxin-negative and *tcdB* PCR-negative and/or if a cultured *C. difficile* isolate did not contain *tcdA* or *tcdB*. Thus, an infant could be classified as co-colonized with both TCD (if stool is *tcdB* PCR-positive) and NTCD (if the single *C. difficile* colony randomly selected from stool culture did not contain *tcdA* or *tcdB* by WGS).

Thirty-two environmental cultures of the general pediatrics clinic were performed after the cohort study was completed. The following surfaces were swabbed with Copan eSwabs (Copan Diagnostics, Murrieta, CA): six surfaces (door knobs, sink/faucet, exam table, parent chairs, otoscope, computer keyboard) in each of four examination rooms; three surfaces (door knobs, sink/faucet, toilet seat) in each of two bathrooms; and two infant weight scales. Liquid transport media from environmental swabs were cultured within 2 hours after collection on taurocholate cycloserine cefoxitin fructose agar and incubated at 37^{o}C for 5 days. Cultures of liquid media were performed under both normal conditions and after heat shock (80^{o}C for 10 minutes).

Genomic Analyses

All cultured *C. difficile* isolates underwent whole genome sequencing using the Illumina MiSeq platform (Illumina, San Diego, CA); DNA extraction, genomic library preparation, sequencing, and assembly bioinformatics analyses have been previously described.[3] *In silico* multilocus sequence typing

(MLST)[4] was performed on assemblies of whole genomic sequences using PubMLST (https://pubmlst.org/cdifficile/), which permitted isolate sequence type (ST) and clade assignment. Isolate genetic relatedness was determined by performing pairwise comparisons of core genome single nucleotide variants (SNVs) among strains, as previously described.[3] Isolates were considered isogenic if they differed by ≤ 2 core genome SNVs (isolates collected <124 days apart) or ≤ 3 core genome SNVs (isolates collected 124-364 days apart), based on analyses of *C. difficile* evolutionary rate in adults by Eyre, et al,[5] which was subsequently validated in infants.[6] Using kSNP (v3.021; https://sourceforge.net/projects/ksnp/), a core genome (core genome definition of SNV loci present in ≥ 95% of isolates) cladogram was generated from the sequenced infant isolates in the present study and 134 previously sequenced clinical isolates that caused CDI at our pediatric medical center between 2011 and 2013.[1, 3, 7] This Whole Genome Shotgun project (Bioproject PRJNA487085) has been deposited at DDBJ/ENA/GenBank under the accession numbers QWUB00000000-QWVR00000000. Individual isolate accession numbers, sequencing information, ST, MLST clade, and toxigenicity are listed in additional Supplementary Material.

Serum Anti-toxin Antibody Assays

Serum samples were analyzed for antibody concentrations to *C. difficile* toxins A and B by enzyme-linked immunosorbent assay (ELISA), as previously described.[8-11] Toxins A and B were purified from the supernatant of a culture of strain VPI 10463 (American Type Culture Collection 43255-FZ, Manassas, VA), as previously described.[12] Non-toxin antigens were prepared from a sonicated whole cell lysate of a NTCD strain (ST-26) isolated from an infant in this cohort. Levels of anti-toxin A and anti-toxin B (IgG, IgA and IgM) and anti-non-toxin antigen (IgG, IgA and IgM) in human sera were measured by direct ELISA using immobilized toxin A, toxin B, or non-toxin antigen by a previously published method.[9] The relative antibody concentration in arbitrary ELISA units for each sample was defined by extrapolation from an internal control standard curve (absorbance versus dilution factor) generated from pooled high titer sera from a different population.

Toxin B Neutralizing Antibody Assay

Serum neutralizing antibody (NAb) titers against toxin B were determined in serum samples. A high throughput NAb assay was performed with modification of a previously described procedure.[13] The assay used an imaging-based phenotypic screen of morphometric features of single cells to identify and titer sera that protected Vero cells from *C. difficile* toxin B-induced cell rounding. Serial two-fold dilutions of serum from 1:20 to 1:10,240 were tested against a standard dose of recombinant[14] VPI10246 toxin B that induced 80% cell rounding of Vero cells (ATCC: CCL81) at 48 hours post-inoculation. De-identified samples were run and imaged in the assay by personnel blinded to patient status. An antibody positive control (ACdTB; Gallus Immunotech, Fergus, Ontario) with neutralizing activity against toxin B was also tested on each plate. Percent cell rounding index (CRI) for each of three to four technical replicates per sample dilution was determined using imaging. Mean CRI was determined at each dilution for each sample or control. The endpoint titer represents the inverse of the dilution of serum that neutralized 50% of toxin activity. Neutralization of 50% of the toxin activity for each sample was calculated by interpolation across the CRI-midpoint determined for each plate. The CRI-midpoint for each plate was the median of two values: the median CRI for wells treated only with toxin (~80-90% CRI) and median CRI for wells treated with medium without toxin. Samples without any neutralizing activity at the lowest dilution (1:20) were assigned a titer of 10 for performance of non-parametric analyses. Thus, a titer of 10 is considered negative for NAb.

Statistics

To account for time required to develop an antibody response, *C. difficile*-colonized infants were grouped based on whether the first colonization event was detected at least 1m prior to serum collection. Proportions were compared between groups using two-tailed Fisher's exact test, and medians were compared between groups using the non-parametric Wilcoxon rank-sum test. We examined the association of TCD colonization with either the highest quartile of ELISA units of IgG against toxins A

and B (compared to the lowest three quartiles), or presence of detectable NAb titers against toxin B, by calculating risk ratios (or risk difference if risk was zero in a group) and using two-tailed Fisher's exact test. Two-sided *P* values < 0.05 were considered statistically significant. Analyses were performed using Stata/IC statistical software, version 12.1 (StataCorp, College Station, TX).

Supplementary Results

Molecular Epidemiology

Infant *C. difficile* colonization molecular epidemiology is illustrated in Figure 2B. Of 57 stools from 23 infants that were positive by GDH EIA and/or *tcdB* PCR, 43 (75%) were culture positive. All isolates were successfully sequenced, assembled, and typed by *in silico* MLST. One novel ST was identified, which had a unique combination of previously identified alleles for each of the 7 MLST housekeeping genes. All isolates belonged to clade 1 (per the PubMLST typing scheme; https://pubmlst.org/cdifficile/),[15] with the exception of the novel ST isolate (subject 14; clade 4). BLAST analysis for *tcdA* and *tcdB* (genes for toxins A and B) identified 26 (60%) TCD strains and 17 (40%) NTCD strains; all TCD strains were positive for both *tcdA* and *tcdB*. All strains were negative for *cdtA*/*cdtB* (genes for binary toxin). This was expected, as MLST did not assign any infant isolates to clades 2 (e.g., epidemic strain BI/NAP1/027) or 5 (e.g., BK/078), which contain binary toxin-producing strains. In total, 8 unique STs were represented among the 26 TCD culture-positive stools (Table S1). Phylogenetic analysis demonstrated that strains colonizing infants clustered with strains previously identified to cause CDI in older children at our pediatric medical center between 2011 and 2013[1, 3, 7] (Figure S1).

Of the 23 infants colonized with TCD and/or NTCD, 16 (70%) had more than one GDH- and/or *tcdB*-positive stool (Figure 2A), indicating persistent *C. difficile* colonization over a several month period; three infants remained colonized with TCD for a period of at least eight months. Ten of these 16 infants (63%)

had multiple positive stool cultures, permitting longitudinal assessment of colonizing *C. difficile* strains. Among the 10 infants with multiple positive stool cultures, six (60%) were colonized with either more than one distinct TCD strain or both TCD and NTCD (Figure 2B). Among the eight infants with multiple isolates belonging to the same ST, there was limited within-host genetic diversity of these isotypic strains over the study period (i.e., 0-1 SNVs between isotypic strains isolated from individual subjects several months apart), suggesting long-term persistence of colonizing strains rather than acquisition of new strains belonging to the same ST. Genetically identical (0-1 SNVs) isolates were identified among two pairs and a trio of unrelated infants, and one pair of siblings (25m difference in age), enrolled in this study (Figure 2B). Because this suggested potential transmission within the clinic (or the home in case of the siblings), environmental cultures of the general pediatrics clinic were obtained from various surfaces in examination rooms, restrooms, and infant scales; all 32 environmental cultures were negative.

Supplementary Table

Table S1: Toxigenic (*tcdA*/*tcdB*-positive) *C. difficile* Sequence Types Isolated from Infant Stool and Frequency of Toxin-EIA Positivity and Detectable Toxin B NAb Titers

Sequence Type (ST)	Number of Stools with ST	Number (%) of Toxin EIA-Positive Stools with ST	Number (%) of Unique Infants with ST (n=16)	Number (%) of Unique Infants with Detectable Nab Titers
ST-2	6	2/6 (33)	3 (19)	0
ST-3	5	4/5 (80)	5 (31)	1/5 (20)
ST-14	8	6/8 (75)	4 (25)	1/4 (25)
ST-42	2	1/2 (50)	2 (13)	1/2 (50)
ST-43	2	2/2 (100)	1 (6)	0
ST-46	1	0/1 (0)	1 (6)	1/1 (100)
ST-49	1	0/1 (0)	1 (6)	0
ST-110	1	1/1 (100)	1 (6)	1/1 (100)

NAb: neutralizing antibody

Supplementary Figures

Figure S1: Core genome cladogram of *C. difficile* isolates colonizing infants between 2014 and 2017 (red font) and those causing CDI in our pediatric medical center between 2011 and 2013 (black font). Infant isolates are labeled with the isolate ID and sequence type (ST) identified by *in silico* MLST. CDI isolates are labeled with the isolate ID, restriction endonuclease analysis group, and ST. Isolates are grouped among 4 of the previously classified *C. difficile* clades (clades 1, 2, 4, and 5 per the MLST scheme; https://pubmlst.org/cdifficile/). NonSp: non-specific REA group. *Non-toxigenic strains (i.e., isolate lacks genes for *tcdA* and *tcdB*).

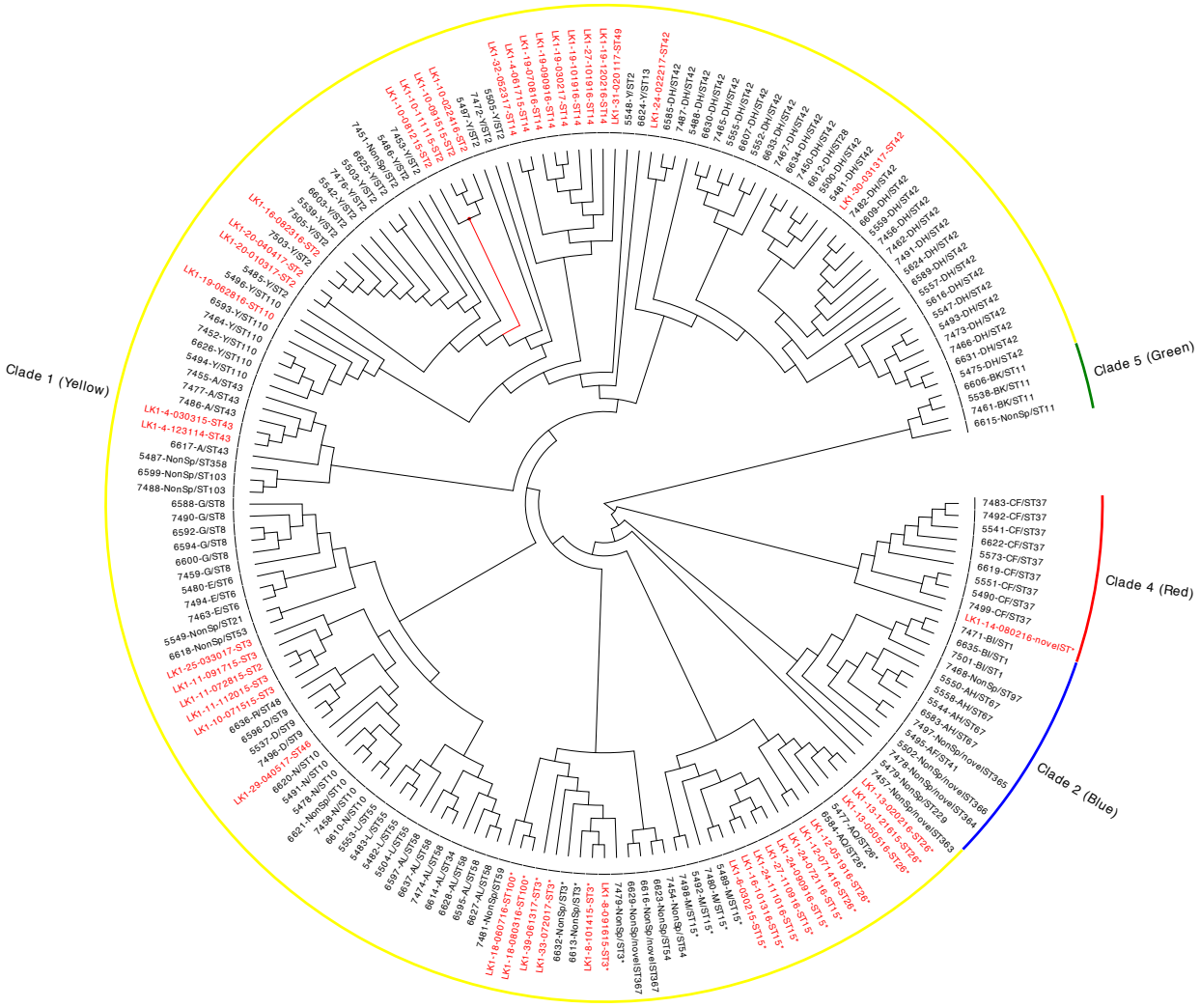

Clade 1 (Yellow)
Clade 5 (Green)
Clade 4 (Red)
Clade 2 (Blue)
7483-CF/ST37
7492-CF/ST37
5541-CF/ST37
6622-CF/ST37
5573-CF/ST37
6619-CF/ST37
5551-CF/ST37
5490-CF/ST37
7499-CF/ST37
LK1-14-080216-novelST-
7471-BI/ST1
6635-BI/ST1
7501-BI/ST1
7468-NonSp/ST97
5550-AH/ST67
5558-AH/ST67
5544-AH/ST67
6583-AH/ST67
7497-NonSp/novelST365
5492-AF/ST41
7478-NonSp/novelST365
5479-NonSp/ST229
7457-NonSp/novelST364
LK1-13-121615-ST26*
LK1-13-050516-ST26*
6594-AO/ST26*
5477-AO/ST26*
LK1-12-051516-ST26*
LK1-12-071416-ST26*
LK1-24-070716-ST15*
LK1-24-090816-ST15*
LK1-27-110916-ST15*
LK1-16-030715-ST15*
5489-M/ST15*
7480-M/ST15*
5492-M/ST15*
7498-M/ST15*
7454-NonSp/ST54
6623-NonSp/novelST867
6616-NonSp/novelST867
6629-NonSp/novelST867
7479-NonSp/ST3*
LK1-8-060916-ST3*
LK1-8-101415-ST3*
6613-NonSp/ST3*
6632-NonSp/ST3*
LK1-33-072017-ST3*
LK1-39-061917-ST3*
LK1-18-080316-ST100*
LK1-18-060716-ST100*
7461-NonSp/ST59
6627-AL/ST58
6595-AL/ST58
6628-AL/ST58
6614-AL/ST58
7474-AL/ST58
6637-AL/ST58
5544-LI/ST55
5482-LI/ST55
5483-LI/ST55
5553-LI/ST55
6610-N/ST10
7456-N/ST10
6621-NonSp/ST10
5491-N/ST10
5478-N/ST10
6620-N/ST10
LK1-29-040517-ST46
7498-D/ST9
5557-D/ST9
6596-D/ST9
6636-R/ST48
LK1-10-071516-ST3
LK1-11-112015-ST3
LK1-11-072815-ST2
LK1-11-091715-ST3
LK1-25-033017-ST3
5549-NonSp/ST21
6618-NonSp/ST53
7463-E/ST6
7494-E/ST6
5480-E/ST6
7459-G/ST8
6600-G/ST8
6594-G/ST8
6592-G/ST8
7490-G/ST8
6588-G/ST8
7488-NonSp/ST103
6599-NonSp/ST103
5487-NonSp/ST358
6617-A/ST43
LK1-4-123114-ST43
LK1-4-030315-ST43
7486-A/ST43
7477-A/ST43
7455-A/ST43
5494-Y/ST110
6626-Y/ST110
7452-Y/ST110
7464-Y/ST110
6593-Y/ST110
LK1-19-062616-ST110
5496-Y/ST110
5485-Y/ST110
LK1-20-040417-ST2
LK1-20-010317-ST2
7503-Y/ST2
LK1-16-082316-ST2
6603-Y/ST2
5542-Y/ST2
7478-Y/ST2
6625-Y/ST2
7451-NonSp/ST2
5486-Y/ST2
7453-Y/ST2
LK1-10-082215-ST2
LK1-10-111115-ST2
LK1-10-022416-ST2
5497-Y/ST2
7472-Y/ST2
5505-Y/ST2
LK1-32-052317-ST14
LK1-4-061717-ST14
LK1-19-070816-ST14
LK1-19-090916-ST14
LK1-19-030217-ST14
LK1-19-101916-ST14
LK1-27-101916-ST14
LK1-19-120216-ST14
LK1-31-020117-ST49
LK1-24-022217-ST42
6624-Y/ST13
5548-Y/ST2
6485-DH/ST42
5487-DH/ST42
5488-DH/ST42
6630-DH/ST42
7465-DH/ST42
5555-DH/ST42
6607-DH/ST42
5552-DH/ST42
6633-DH/ST42
7467-DH/ST42
6634-DH/ST42
7450-DH/ST42
6612-DH/ST42
5500-DH/ST28
5481-DH/ST42
LK1-30-031317-ST42
7482-DH/ST42
6609-DH/ST42
5559-DH/ST42
7456-DH/ST42
7462-DH/ST42
7491-DH/ST42
5824-DH/ST42
6589-DH/ST42
5557-DH/ST42
5616-DH/ST42
5547-DH/ST42
5493-DH/ST42
7473-DH/ST42
7466-DH/ST42
6631-DH/ST42
5475-DH/ST42
6606-BK/ST11
5538-BK/ST11
7461-BK/ST11
6615-NonSp/ST11

Figure S2. Association between toxigenic *C. difficile* (TCD) colonization with toxin EIA positivity and serum IgA and IgG against toxins A and B. Data expressed as arbitrary ELISA units on a natural logarithmic scale. Boxes delineate the median and IQR, whiskers delineate the upper and lower adjacent values (within 1.5 x IQR), and isolated data points are outliers. Non-colonized infants were compared to both toxin EIA-positive and negative infants with TCD colonization. Bolded *p* values are statistically significant.

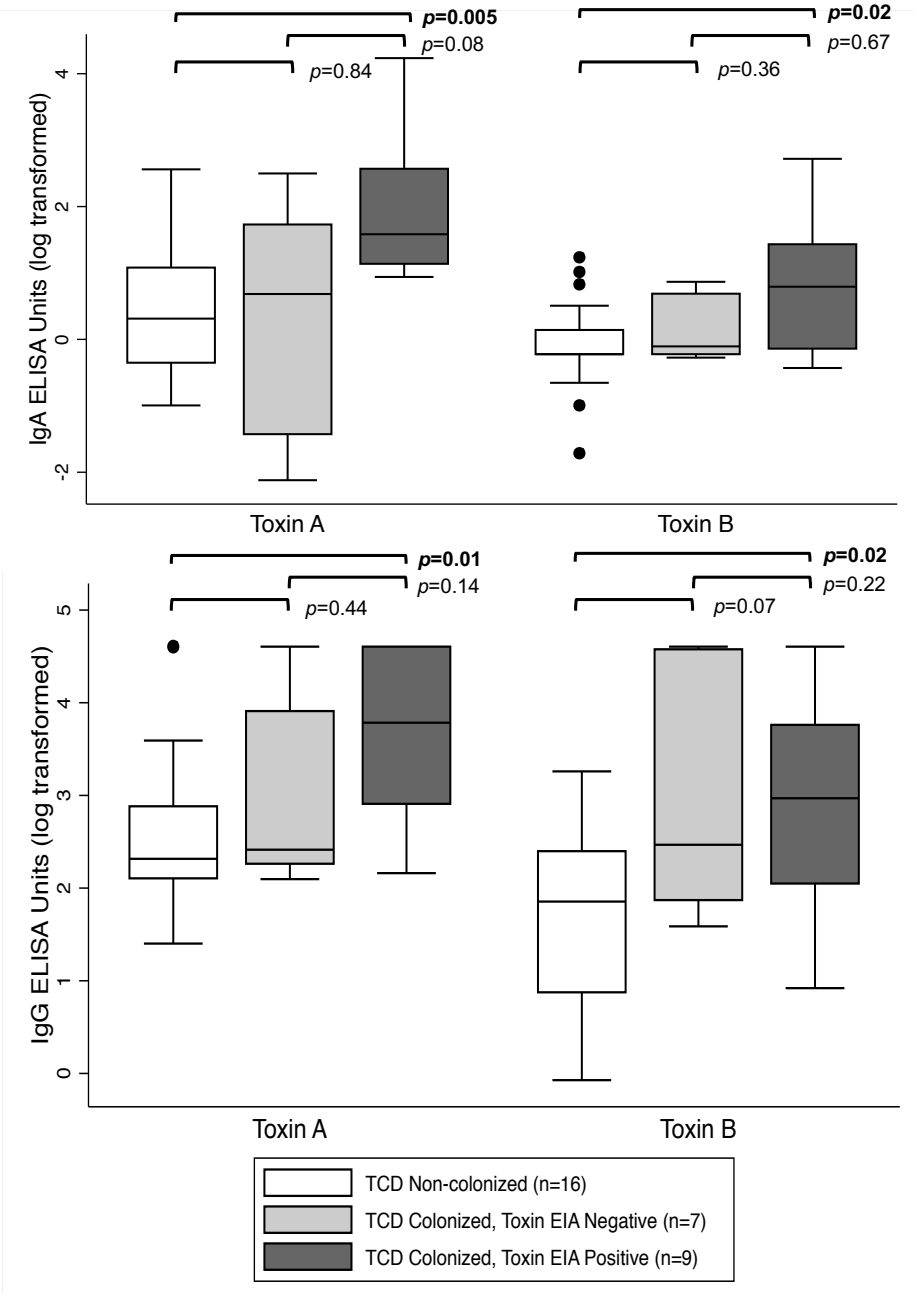


Figure S3. Association between toxigenic *C. difficile* (TCD) colonization with toxin EIA positivity and toxin B NAb titers. Data expressed on a natural logarithmic scale. Boxes delineate the median and IQR, whiskers delineate the upper and lower adjacent values (within 1.5 x IQR), and isolated data points are outliers. Non-colonized infants were compared to both toxin EIA-positive and negative infants with TCD colonization. Bolded *p* values are statistically significant.

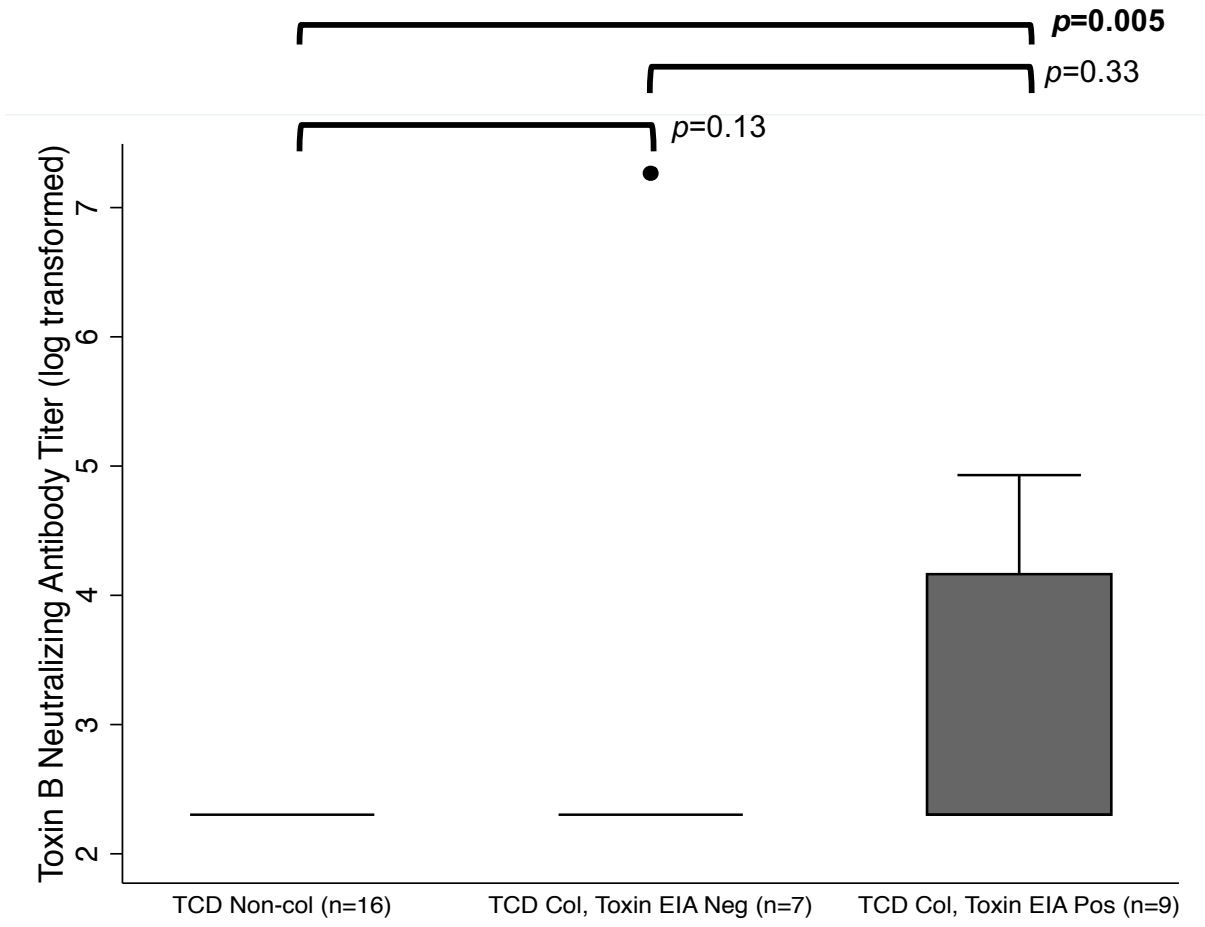


Figure S4. Comparisons of IgG against toxins A and B in cord blood (from unrelated mothers), infants without toxigenic *C. difficile* (TCD) colonization, and infants with TCD colonization for at least 1m duration. Data expressed as arbitrary ELISA units on a natural logarithmic scale. Boxes delineate the median and IQR, whiskers delineate the upper and lower adjacent values (within 1.5 x IQR), and isolated data points are outliers. Cord blood log transformed ELISA units for IgG against toxins A and B were compared to both non-colonized infants and infants with TCD colonization for at least 1m duration. Bolded *p* values are statistically significant.

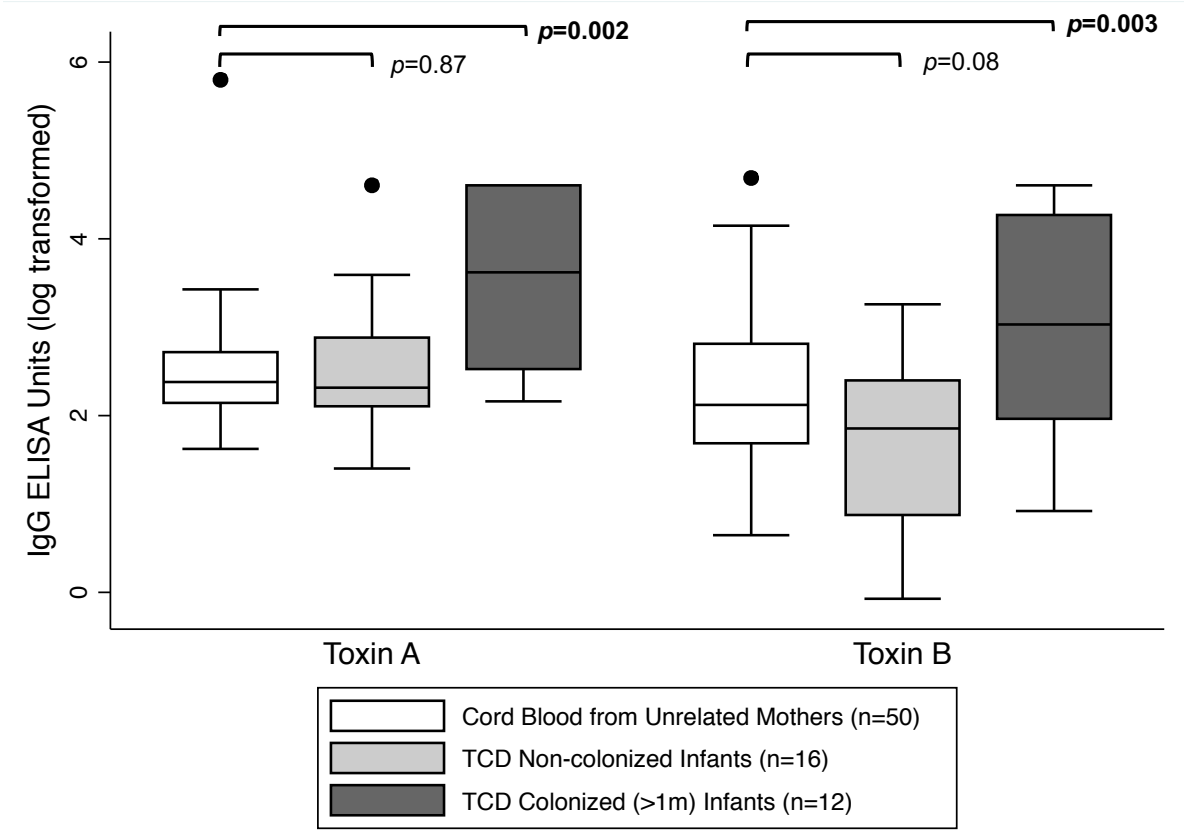
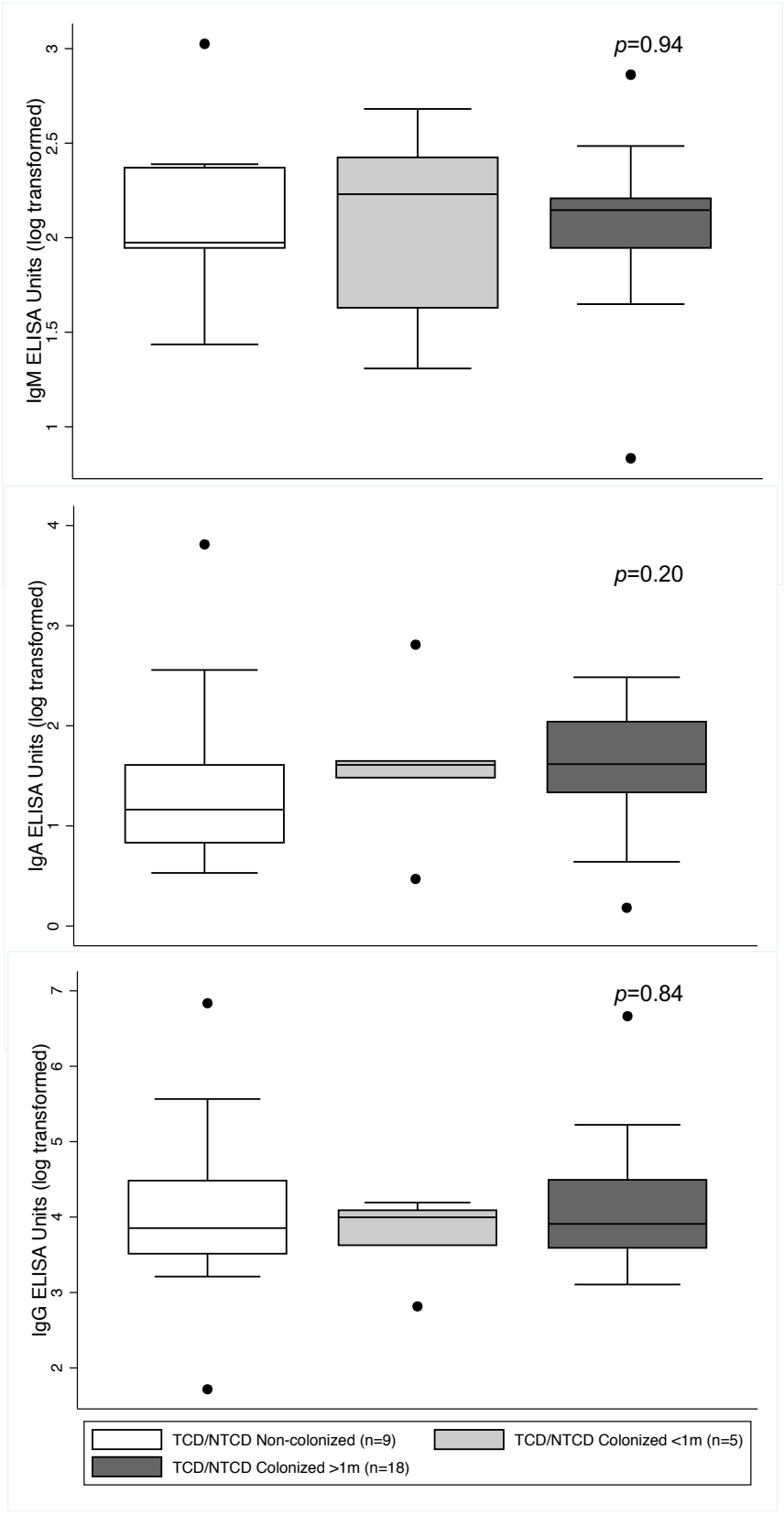


Figure S5. Association between toxigenic (TCD) and/or non-toxigenic (NTCD) *C. difficile* colonization and serum IgM, IgA, and IgG against non-toxin antigens. Data expressed as arbitrary ELISA units on a natural logarithmic scale. Boxes delineate the median and IQR, whiskers delineate the upper and lower adjacent values (within 1.5 x IQR), and isolated data points are outliers. Reported *p* values are for comparisons between non-colonized infants and infants with TCD and/or NTCD colonization for at least 1m duration.


IgM ELISA Units (log transformed)
1
1.5
2
2.5
3
p=0.94
IgA ELISA Units (log transformed)
0
1
2
3
4
p=0.20
IgG ELISA Units (log transformed)
2
3
4
5
6
7
p=0.84
TCD/NTCD Non-colonized (n=9)
TCD/NTCD Colonized <1m (n=5)
TCD/NTCD Colonized >1m (n=18)

Supplementary References

1. Kociolek LK, Patel SJ, Shulman ST, Gerding D. Molecular epidemiology of Clostridium difficile infections in children: A retrospective cohort study. Infect Control Hosp Epidemiol **2015**; 36(4): 445-51.
2. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol **1990**; 215(3): 403-10.
3. Kociolek LK, Gerding DN, Espinosa RO, Patel SJ, Shulman ST, Ozer EA. Clostridium difficile Whole Genome Sequencing Reveals Limited Transmission Among Symptomatic Children: A Single-Center Analysis. Clin Infect Dis **2018**; 67(2): 229-34.
4. Griffiths D, Fawley W, Kachrimanidou M, et al. Multilocus sequence typing of Clostridium difficile. J Clin Microbiol **2010**; 48(3): 770-8.
5. Eyre DW, Cule ML, Wilson DJ, et al. Diverse sources of C. difficile infection identified on whole-genome sequencing. N Engl J Med **2013**; 369(13): 1195-205.
6. Stoesser N, Eyre DW, Quan TP, et al. Epidemiology of Clostridium difficile in infants in Oxfordshire, UK: Risk factors for colonization and carriage, and genetic overlap with regional C. difficile infection strains. PloS One **2017**; 12(8): e0182307.
7. Kociolek LK, Gerding DN, Hecht DW, Ozer EA. Comparative genomics analysis of Clostridium difficile epidemic strain DH/NAP11/106. Microbes Infect **2018**; 20(4): 245-53.
8. Hourigan SK, Chirumamilla SR, Ross T, et al. Clostridium difficile carriage and serum antitoxin responses in children with inflammatory bowel disease. Inflammatory Bowel Dis **2013**; 19(13): 2744-52.

9. Kyne L, Warny M, Qamar A, Kelly CP. Asymptomatic carriage of Clostridium difficile and serum levels of IgG antibody against toxin A. N Engl J Med **2000**; 342(6): 390-7.

10. Kyne L, Warny M, Qamar A, Kelly CP. Association between antibody response to toxin A and protection against recurrent Clostridium difficile diarrhoea. The Lancet **2001**; 357(9251): 189-93.

11. Hughes M, Qazi T, Berg A, et al. Host Immune Response to Clostridium difficile Infection in Inflammatory Bowel Disease Patients. Inflammatory Bowel Dis **2016**; 22(4): 853-61.

12. Sullivan NM, Pellett S, Wilkins TD. Purification and characterization of toxins A and B of Clostridium difficile. Infect Immun **1982**; 35(3): 1032-40.

13. Xie J, Zorman J, Indrawati L, et al. Development and optimization of a novel assay to measure neutralizing antibodies against Clostridium difficile toxins. Clin Vaccine Immunol **2013**; 20(4): 517-25.

14. Yang G, Zhou B, Wang J, et al. Expression of recombinant Clostridium difficile toxin A and B in Bacillus megaterium. BMC Microbiol **2008**; 8: 192.

15. Knight DR, Elliott B, Chang BJ, Perkins TT, Riley TV. Diversity and Evolution in the Genome of Clostridium difficile. Clin Microbiol Rev **2015**; 28(3): 721-41.